УДК 519.21.

ДОВІРЧІ ГРАНИЦІ ДЛЯ НЕВІДОМОЇ ДИСПЕРСІЇ У ВИПАДКУ НЕВІДОМОГО МАТЕМАТИЧНОГО СПОДІВАННЯ.

Тупко Н.П., Васильєва Н.С. (Одеська державна академія будівництва та архітектури, м.Одеса)

На базі оцінок для дисперсії у випадку невідомого математичного сподівання [3] побудовано довірчі границі для невідомої дисперсії за допомогою правила 3σ.

Розв`язок задач, пов`язаних з надійністю продукції та вибором оптимальних технологічних варіантів при виробництві будівельних матеріалів, у сучасних умовах є дуже важливою проблемою. Це пов`язано як з постійним ускладненням самої технології виготовлення будматеріалів, так і з необхідністю поліпшення економічних показників виробництва. Для розв`язання технологічних задач використовують імовірностно-статистичні методи. Застосування цих методів дозволяє робити статистичний аналіз накопичених даних про технологічні процеси, рецептури та властивості сировини і матеріалів з метою узагальнення інформації у вигляді формул. Одна із головних задач, які при цьому виникають, це визначення по експериментальним даним закону розподілу випадкових величин і відповідних числових характеристик.

До основних характеристик випадкових величин відносяться математичне сподівання та дисперсія. Математичне сподівання представляє характеристику розташування можливих значень випадкової величини, а дисперсія є мірою розташування цих значень відносно математичного сподівання. Для оцінки невідомого математичного сподівання побудовані численні теорії і отримані фундаментальні результати . Але проблемі оцінки невідомої дисперсії приділяли значно менше уваги.

У роботі запропановано побудова довірчих границь для невідомої

дисперсії у випадку невідомого математичного сподівання на базі оцінок отриманих у попередніх працях.

П.1. Зміщена та незміщена оцінки невідомої дисперсії у випадку невідомого математичного сподівання.

Для дисперсії розглянемо дві оцінки незміщену і зміщену відповідно:

$$\widetilde{S}^2 = \frac{1}{n-1}\sum_{k=1}^{n}(x_k - \bar{x})^2 \quad (1)$$

і

$$S_*^2 = \frac{1}{n}\sum_{k=1}^{n} x_k^2 - \bar{x}^2 \quad (2)$$

Математичне сподівання яких дорівнює:

$$m(\widetilde{S}^2) = \sigma^2,\ m(S_*^2) = \frac{n-1}{n}\sigma^2 \quad (3)$$

Середнє квадратичне відхилення цих оцінок, згідно результатам отриманим у роботі [3], відповідно :

$$\delta^2(\widetilde{S}^2) = D(\widetilde{S}^2) = \frac{1}{n}m(x^4) + \frac{(n-1)^2+2}{n(n-1)}(m^2+\sigma^2)^2 + \frac{4-2(n-1)}{n(n-1)}(n-2)m^2(m^2+\sigma^2) -$$
$$-\frac{4}{n}m(x^3)m + \frac{(n-2)(n-3)}{n(n-1)}m^4 - \sigma^4 = \frac{1}{n}m(x^4) - \frac{4}{n}m(x^3)m + m^4\frac{3}{n} + m^2\sigma^2\frac{6}{n} +$$
$$+\sigma^4\frac{3-n}{n(n-1)} \quad (4)$$

і

$$\delta^2(S_*^2) = \frac{(n-1)^2}{n^2}\left[\frac{1}{n}m(x^4) + \frac{(n-1)^2+2}{n(n-1)}(m^2+\sigma^2)^2 + \frac{4-2(n-1)}{n(n-1)}(n-2)m^2 \times\right.$$
$$\left.\times(m^2+\sigma^2) - \frac{4}{n}m(x^3)m + \frac{(n-2)(n-3)}{n(n-1)}m^4\right] + \frac{2-n}{n}\sigma^4 = \frac{(n-1)^2}{n^2}\left[\frac{1}{n}m(x^4) -\right.$$
$$\left.-\frac{4}{n}m(x^3)m + m^4\frac{3}{n} + m^2\sigma^2\frac{6}{n}\right] + \sigma^4\frac{5n-n^2-3}{n^3}, \quad (5)$$

де σ^2 - дисперсія, m - математичне сподівання, $m(x^k)$ - момент k -го порядка.

У праці [3] доведено, що для нормально та експоненційно розподілених випадкових величин $\delta^2(S_*^2) < \delta^2(\widetilde{S}^2)$, тобто зміщена оцінка S_*^2 є більш точною ніж незміщена $\widetilde{S}^2$, принаймні для розглянутих розподілів.

П.2. Довірчі границі для невідомої дисперсії у випадку невідомого маткматичного сподівання.

Якщо випадкова величина x нормально розроділена, то згідно [5] для випадкової величини справедливо правило трьох сігм , тобто:

$P\{-3\sigma + x < m(x) < 3\sigma + x\} > 0{,}95$, де $\sigma = \sqrt{D(x)}$ -середньо квадратичне відхилення.

По аналогії з цим правилом побудуємо довірчі границі для дисперсії довільного розподілу на базі оцінок $S_*^{\,2}$ та $\widetilde{S}^{2}$ і підрахуємо рівень значущості.

Враховуючи , що $m(\widetilde{S}^{2}) = \sigma^{2}$, $m(S_*^2) = \dfrac{n-1}{n}\sigma^2$ маємо наступні співвідношення:

$$-3\sigma(\widetilde{S}^{2}) + \widetilde{S}^{2} < \sigma^{2} < 3\sigma(\widetilde{S}^{2}) + \widetilde{S}^{2}, \qquad (6)$$

$$-\frac{3n}{n-1}\delta(S_*^{\,2}) + \frac{n}{n-1}S_*^{\,2} < \sigma^2 < \frac{3n}{n-1}\delta(S_*^{\,2}) + \frac{n}{n-1}S_*^{\,2}. \qquad (7)$$

І відповідно інтервали:

$$\Upsilon_1 = (-3\sigma(\widetilde{S}^{2}) + \widetilde{S}^{2}, 3\sigma(\widetilde{S}^{2}) + \widetilde{S}^{2}), \qquad (8)$$

$$\Upsilon_2 = \left(-\frac{3n}{n-1}\delta(S_*^{\,2}) + \frac{n}{n-1}S_*^{\,2}, \frac{3n}{n-1}\delta(S_*^{\,2}) + \frac{n}{n-1}S_*^{\,2}\right). \qquad (9)$$

Нехай $x_1, x_2, \ldots, x_n$ -вибірка можливих незалежних значень випадкової величини x, тоді враховуючи (1)-(5) і використовуючи замість центрального моменту k -го порядку $L_k = m(x^k)$ оцінку $\frac{1}{n}\sum_{i=1}^{n} x_i^k$, завжди можна побудувати довірчі інтервали (8) або (9), для невідомої дисперсії σ^2, де

$$\delta^2(\widetilde{S}^2) = \frac{1}{n^2}\sum_{i=1}^{n} x_i^4 - \frac{4}{n^3}\sum_{i=1}^{n} x_i^3 \sum_{i=1}^{n} x_i + \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^4 \frac{3}{n} + \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2 \left(\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} x_i^k\right)^2\right)\frac{6}{n} +$$
$$+\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} x_i^k\right)^2\right)^2 \frac{3-n}{n(n-1)}$$

$$\delta^2(S_*^{\,2}) = \frac{(n-1)^2}{n^2}\left[\frac{1}{n^2}\sum_{i=1}^{n} x_i^4 - \frac{4}{n^3}\sum_{i=1}^{n} x_i^3 \sum_{i=1}^{n} x_i + \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^4 \frac{3}{n} +\right.$$
$$\left. + \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2 \left(\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} x_i^k\right)^2\right)\frac{6}{n}\right] + \left(\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} x_i^k\right)^2\right)^2 \frac{5n - n^2 - 3^2}{n^3}$$

П.3. Розрахунок рівнів значущості довірчих інтервалів для невідомої дисперсії у випадку нормального та експоненціального розподілів методами машинного моделювання.

За допомогою програмної реалізації на мові вищого рівня Turbo Pascal 7.0, для нормального та експоненціального розподілів, будуються довірчі інтервали (8) на основі вибірок розмірності N і підраховуються рівні значущості для кожного з розподілів: як відношення кількості інтервалів (k), які задовільняюь (6), до загальної кількості (m) побудованих інтервалів. Аналогічно підраховується рівень значущості для довірчих інтервалів (9).

Маємо наступні результати.

1. Нормальний розподіл N(0,1).

m-загальна кількість довірчих інтервалів;
N-розмір вибірки.

Таблиця 1

Результати підрахунків рівнів значущості у випадку нормально розподіленої величини.

m	N	Рівень значущості для Y_1	Рівень значущості для Y_2
5	10	6.00000000000E-01	1.00000000000E+00
5	20	8.00000000000E-01	1.00000000000E+00
5	40	1.00000000000E-01	1.00000000000E+00
5	70	0.00000000000E+00	1.00000000000E+00
10	10	5.00000000000E-01	1.00000000000E+00
10	20	5.00000000000E-01	1.00000000000E+00
30	10	2.33333333333E-01	1.00000000000E+00
30	20	5.77777777777E-01	1.00000000000E+00
30	40	9.00000000000E-01	1.00000000000E+00
50	10	2.60000000000E-01	1.00000000000E+00
50	20	5.80000000000E-01	1.00000000000E+001.00
100	10	3.00000000000E-01	000000000E+001.00000
100	20	6.00000000000E-01	000000E+001.00000000
100	40	7.50000000000E-01	000E+00
100	100	0.00000000000E+00	1.00000000000E+00

2. Експоненціальний розподіл з параметром λ =1.

m-загальна кількість довірчих інтервалів;
N-розмір вибірки.

Таблиця 2

Результати підрахунків рівнів значущості у випадку експоненціально розподіленої величини.

m	N	Рівень значущості для Y_1	Рівень значущості для Y_2
5	20	0.0000000000E+00	1.0000000000E+00
5	40	4.0000000000E-01	1.0000000000E+00
5	70	4.0000000000E-01	1.0000000000E+00
10	10	0.0000000000E+00	0.0000000000E+00
10	20	0.0000000000E+00	1.0000000000E+00
30	10	0.0000000000E+00	0.0000000000E+00
30	20	0.0000000000E+00	1.0000000000E+00
30	40	3.3333333333E-02	1.0000000000E+00
50	10	0.0000000000E+00	0.0000000000E+00
50	20	0.0000000000E+00	1.0000000000E+00
50	40	1.2000000000E-01	1.0000000000E+00
100	10	0.0000000000E+00	0.0000000000E+00
100	20	0.0000000000E+00	9.9000000000E-01
100	40	8.0000000000E-02	1.0000000000E+00
100	100	0.0000000000E+00	1.0000000000E+00

Висновок. Одержані у роботі експериментальні результати п.3 підтверджують теоретичні результати викладені у п.1,тобто для нормального та експоненціального розподілів оцінка $S_*^{\,2}$ є кращою ніж $\widetilde{S}^2$.

Литература:

1. Г.И.Байдак, М.Ш.Браверман, Ю.И.Петунин. Аддитивность дисперсии характеристическое свойство гильбертового пространства $L_2(\Omega, A, \mu)$.//Функциональный анализ и его приложения-1983.-17, вып.3.-С.66-68.

2. Ю.И.Петунин. Приложение теории случайных процессов в биологии и медицине -К: Наукова думка,1981. -320 с.

3.Петунин Ю.И., Тупко Н.П. Теория квадратичных оценок дисперсии. // Український математичний журнал.-1999.-Том 51 № 9.-С.1217-1231.

4. Ю.Г.Курицын, Ю.И.Петунин. К теории линейных оценок математического ожидания случайного процесса// Теория вероятности и математическая статистика.-1970.-Вып.3.-С.80 - 92.

5. Ван дер Варден Б.Л. Математическая статистика.-М.:Изд-во иностр. Лит.,1960.- 436 с.

6. Гарольд Крамер. Математические методы статистики.-М.:Мир,1975.-648с.

7. В.Вознесенский. Статистические решения в технологических задачах.-Кишенёв.:Картя Молдовеняске,1969.-232с.

8. В.И.Романовский. Математическая статистика.-Ташкент: Изд-во АН УзССР,1961.-638 с.